Exploratory and Confirmatory Factor Analysis in Gifted Education: Examples With Self-Concept Data

Jonathan A. Plucker

Factor analysis allows researchers to conduct exploratory analyses of latent variables, reduce data in large datasets, and test specific models. The purpose of this paper is to review common uses of factor analysis, provide general guidelines for best practices, illustrate these guidelines with examples using previously published self-concept data, and discuss common pitfalls and ways to avoid them.

Introduction

Factor analysis is among the most versatile and controversial techniques for analyzing data in the behavioral and social sciences. Factor analysis is commonly used to analyze complex data sets within the field of gifted education, yet it is often misused and misinterpreted. For example, Gould's 1981 description of factor analysis is a popular treatment of the topic, yet Carroll (1995) criticized Gould's interpretation of factor analysis. This commentary introduces readers to general issues surrounding factor analysis and suggests some best practices when using and reporting results of factor analyses in gifted education. Interested readers should also consult technical treatments of the topic that provide step-by-step guidance, such as those provided by Pedhazur and Schmelkin (1991), Tabachnick and Fidell (2001), Hurley et al. (1997), Kieffer (1999), and Byrne (1998, 2001), among many others.

Factor Analysis

Factor analysis is most often used to provide evidence of construct validity for an instrument or assessment. For example, con-

---

Jonathan A. Plucker is Associate Professor of Learning Cognition and Instruction and Associate Professor of Cognitive Science at Indiana University, Bloomington.

Journal for the Education of the Gifted. Vol. 27, No. 1, 2003, pp. 20–35.

sider the case of a researcher who has designed a self-concept scale that produces two scores: an academic self-concept score and a nonacademic self-concept score. The researcher uses factor analysis to analyze data collected with the instrument to determine the number of factors that can be extracted from the data. If the analyses provide evidence of one, three, or more factors, the evidence of construct validity will be weak. If the analyses provide evidence of the existence of two factors among the collected data, the researcher will have gathered evidence of construct validity.

Factor analysis is controversial because, in at least one of its forms, it is quite subjective when compared to other statistical techniques. Wide-ranging opinions exist about many of these subjective issues, providing the interested scholar with several confusing—and often contradictory—lists of suggestions for conducting factor analyses of data. This paper uses examples from a previously published self-concept study to illustrate best practices in the use of factor analysis.

Statistically, most approaches to factor analysis involve the investigation of correlations among scores on several variables in an attempt to see how the variable scores "clump together." For example, to return to the self-concept example, the researcher hopes to find that the variables representing academic self-concept correlate highly with each other, but poorly with nonacademic self-concept items. In the same vein, the researcher hopes that nonacademic self-concept items correlate more highly with each other than with academic self-concept items. If all of the variables correlate with all other variables, evidence exists for one factor, but not two relatively independent factors.

In the following sections, the examples are drawn from self-concept data collected from adolescents participating in the Duke University Talent Identification Program (TIP). These students completed the long form of the Self-Description Questionnaire II (SDQII; Marsh, 1992), a popular measure of adolescent self-concept. The SDQII includes 102 items indicating levels of self-concept in 11 dimensions. The examples in this paper include only the 10 item pairs representing the math and verbal self-concept scales. Factor analysis will be used to examine the factor structure of these two scales (i.e., do the items from the two scales appear to measure two distinct aspects of self-concept, or do they measure one aspect that can be interpreted as general academic self-concept?).[1] In the examples provided in this paper, the data is drawn from 339 adolescents participating in the TIP during the summer of 1995.

Exploratory Factor Analysis (EFA)

Researchers use exploratory factor analysis when they are interested in (a) attempting to reduce the amount of data to be used in subsequent analyses or (b) determining the number and character of underlying (or latent) factors in a data set. Although these purposes sound very similar, they are slightly different and lead to different statistical approaches. When attempting to reduce data, statisticians often recommend the use of principal components analysis, although factor analysis can also be used to perform this function. In most cases, factor analysis is recommended when attempting to determine the presence of latent factors within a set of variables.

The term *exploratory* is used for good reason: EFA does not test a model of factor structure. Rather, the computer program explores the data set in search for statistically justified factors. This leads to the subjectivity mentioned earlier: Determining how many factors to select is a subjective and often arbitrary process. One set of factors may be interpreted very differently by different researchers. Later in this article, we suggest methods for increasing objectivity in the selection and interpretation of factors, but at this point it is only important to understand that EFA is truly an exploratory process that does not formally examine the validity of a priori theory.

Most major statistical computer programs, including SPSS and SAS, allow for EFA. Each program provides different options and techniques (see Tabachnick & Fidell, 2001, pp. 649–651), although the basic steps are similar: extraction, selection, rotation, and interpretation.

The first step in any EFA is the *extraction of factors*, during which the computer program examines the covariance among the numerous variables in an attempt to identify factors underlying the data. Tabachnick and Fidell (2001) noted that principal components analysis (PCA) and principal factors are the most commonly used extraction techniques, but researchers regularly utilize several other techniques, including maximum likelihood and alpha factor extraction. Indeed, Pedhazur and Schmelkin (1991) argued against the use of PCA as an EFA extraction technique, contending that PCA and EFA are different statistical techniques with different purposes.[2] Each extraction technique derives factors in a particular way. Solutions produced by each extraction technique will vary. These differences are often small (especially in large data sets), but researchers need to be aware that the same data set may produce different results if two different extraction techniques are used. Maximum likelihood extraction results for the self-concept data are presented in Table 1.

Table 1

Initial Eigenvalues and Results for Maximum Likelihood Extraction of Academic Self-Concept Data

Factor	Eigenvalue	Parallel analysis eigenvalue	% Variance	Cumulative % variance	% Variance after extraction	% Variance after rotation
1	3.81	1.27	38.1	38.1	33.6	30.2
2	2.97	1.19	29.7	67.7	27.9	32.5
3	1.04	1.17	10.4	78.2		
4	.45	1.08	4.5	82.7		
5	.41	1.01	4.1	86.8		
6	.37	.96	3.7	90.6		
7	.31	.93	3.1	93.7		
8	.25	.88	2.4	96.1		
9	.21	.76	2.1	98.2		
10	.18	.75	1.8	100.0		

Factor selection is the most controversial aspect of EFA, primarily due to the numerous and wide-ranging strategies for determining the number of factors. Regardless of the exact technique employed, extraction provides the researcher with several pieces of helpful information. The most important of these is the eigenvalue of each extracted factor. The eigenvalue is a measure of variance, with a value greater than 1.0 often interpreted as being meaningful. Eigenvalues can also be plotted against factors to perform the scree test as an aid during factor selection. Parallel analysis is an interesting strategy that requires factor analysis of a similar data set composed of random numbers. If the eigenvalue for the first factor using real data exceeds the eigenvalue for the first factor using random data, then the factor should be selected (see instructions for parallel analysis in Thompson & Daniel, 1996). Additional strategies for selecting the number of factors are discussed in the overview provided by Thompson and Daniel.

Factor selection for the self-concept data was based on the data in Table 1 and the scree plot depicted in Figure 1. The traditional guide for selecting factors with eigenvalues of 1 or greater suggests three factors, interpretation of the scree plot suggests three factors,

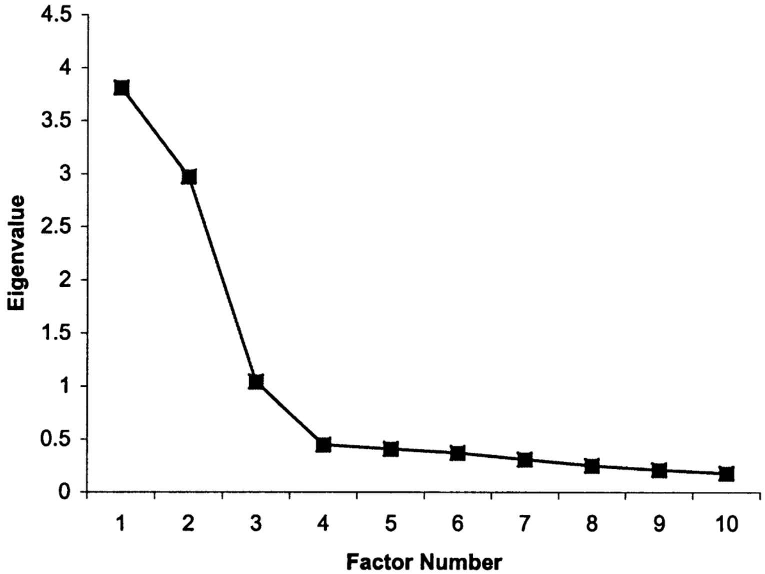

Figure 1. Scree plot for extracted factor eigenvalues for academic self-concept.

and parallel analysis suggests the presence of two factors. Given that theoretical considerations support the existence of two factors, and the observation that a third factor would consist of only one variable, a two-factor solution was chosen for use in this study.[3]

After extraction and factor selection, researchers can interpret the results, or they can *rotate the factors* to produce a better fit between the data and factors. The primary purpose of rotation is to make the results easier to interpret. There are many ways to rotate factors, but they are all either *orthogonal* (rotating factors so that they are not correlated with each other) or *oblique* (allowing rotated factors to correlate). Rotating factors generally produces a better fit for the factors to the data, increasing the ease of interpretation.

For the self-concept data, I rotated the factor obliquely, based on the belief that mathematical and verbal achievement test scores are usually highly correlated. The correlation between the rotated factors is -.16, suggesting that the two factors are, at most, weakly and negatively correlated.

The final step in EFA is the *interpretation of factors*, which is dependent on interpretation of two matrices. The pattern matrix represents the relationship between each variable and factor, controlling for other factors. The structure matrix represents the correlation

Table 2

Pattern and Structure Matrix Loadings and Communalities for Academic Self-Concept Items

Item	Pattern matrix: Math SC factor	Pattern matrix: Verbal SC factor	Structure matrix: Math SC factor	Structure matrix: Verbal SC factor	h^2
Math 1	.80	-.09	.81	-.21	.67
Math 2	.83	.03	.83	-.10	.68
Math 3	-.12	.03	-.12	.05	.02
Math 4	.88	.05	.88	-.09	.77
Math 5	.90	.15	.88	.01	.79
Verbal 1	-.08	.75	-.19	.76	.59
Verbal 2	.01	.84	-.12	.84	.70
Verbal 3	-.11	.86	-.25	.88	.79
Verbal 4	.14	.79	.02	.77	.61
Verbal 5	.00	.73	-.11	.73	.53

between each variable and each factor, without control for the variables' correlations with other factors. For orthogonal rotations, the pattern and structure matrices are identical. For oblique rotations, however, these matrices may produce very different results; and Pedhazur and Schmelkin (1991) recommend interpreting the pattern matrix in this situation. A range of standards are used to determine whether a factor loading in a pattern matrix is practically significant. Many researchers use a cutoff of .30, others use .35, and some use .40 or higher. In the end, the researcher needs to consider ease of factor interpretation when setting a cutoff for loading interpretation.

Statistical analysis programs also provide an estimate of the communality, the amount of variance for each of the observed variables accounted for by the factor solution. High values indicate factor solutions that explain a sizeable degree of variance for a particular variable or set of variables; low values suggest a poor fit between the observed variables and factor solution. *When the factors are independent* (i.e., if the factors are not rotated or are rotated orthogonally), these estimates can be calculated by summing the squared loadings of each item on each factor. In Table 2, communalities are not equal to the sum of the squared loadings due to the correlation of the factors (i.e., oblique rotation).

The communalities and pattern and structure matrices for the self-concept data are presented in Table 2. Loadings in both matrices suggest that most math items load highly on the math factor and poorly on the verbal factor, and all verbal items load highly on the verbal factor and poorly on the math factor. The third math item loads poorly on both factors, resulting in a low-communality estimate and suggesting that the construct validity of this item is not well supported.

Feldhusen, Dai, and Clinkenbeard (2000) used EFA to examine the underlying factor structure of 176 students' scores on the 28-item, author-created Cooperation/Competition Scale. Four factors emerged from the data: Cooperation, Competition-Outcome (desire to win or outperform others), Competition-Process (enjoyment of competition as a mode of learning), and Disengagement (preference for withdrawing from cooperative or competitive situations). The authors used varimax rotation (extraction information was not included in the article). The EFA provided evidence that competitive learning situations can motivate gifted students both in outcome-oriented (often associated with negative outcomes) and process-oriented (reflecting constructs often associated with positive achievement and affective outcomes) ways. This conclusion suggests that instructional design based on unidimensional models of competition (or a lack thereof) may be simplistic and inefficient.

In another example, Masten, Morse, and Wenglar (1995) used EFA to investigate the factor structure of a popular intelligence test with a sample of Mexican American students referred for identification screening for a gifted program. Researchers often perform these analyses to see if previously published factor structures for instruments are the same for different samples of students. Scores on the instrument in this study are often associated with a three-factor EFA model. Masten et al. also found evidence of three factors with their sample, but one factor was different in scope from the traditional three-factor model. Masten et al. employed maximum likelihood extraction with orthogonal (varimax) rotation of factors. The authors concluded that "these findings may suggest different interpretations of [intelligence test] scores for [Mexican American students] referred for gifted programs and a reexamination of cut-off scores for admission" (p. 131) to programs for the intellectually gifted.

Confirmatory Factor Analysis (CFA)

As the title of CFA suggests, the major distinction between confirmatory and exploratory factor analysis is the ability (and necessity)

to test a specific model of factor structure when using CFA. This allows for models in which not all variables are correlated with all factors. Furthermore, CFA provides researchers with the ability to correlate errors and test whether a specific model is equivalent across data from distinct groups. Several structural equation modeling programs can be used to perform CFA, with LISREL and AMOS being the most popular. Byrne has written two accessible books on how to use both programs: Her 1998 text addresses applications of LISREL, and her 2001 book covers the use of AMOS.

The availability of these computer programs, especially AMOS with its graphical interface, has streamlined the use of CFA. Personally, I find the steps for conducting CFA to be more straightforward and less subjective than those for EFA. In general, CFA requires five steps: model specification, model estimation (fitting the model), evaluation of fit, model modification, and interpretation of loadings and related statistics.

In analyzing the self-concept data, we *specified several models*: (a) the independence model, which hypothesized no relationship among any of the 10 variables and served as a baseline comparison to subsequent models; (b) the saturated model, a "best case" model in which every variable is correlated to every other variable; (c) a one-factor model that represents the hypothesis that the 10 variables assess one underlying construct; (d) a two-factor model that posits the existence of verbal and mathematical self-concept factors; (e) and a two-factor model that proposes the existence of correlated verbal and mathematical self-concept factors.

After specifying the five models, they were *fit to the data*. AMOS provides extensive written output for each estimated model, with a wide variety of statistics to aid in determining how well each model fits the data (Table 3). All of the fit statistics have strengths and weaknesses, and selection of appropriate goodness-of-fit measures is often a matter of personal preference. Some of the most commonly recommended statistics for *evaluation of fit* are provided in Table 4 for the tested self-concept models. Fit statistics suggest that the two-factor models are better fitting than the one-factor and independence models, although the results are mixed when comparing the two-factor and two-correlated-factor models. Given the low correlation (-.18) in the correlated model and a desire for parsimony, the two-factor model is chosen for modification and interpretation.

CFA programs often provide *modification indexes* to help researchers determine if adding parameters to a specific model would increase the goodness-of-fit. In the self-concept example, AMOS suggested correlating several of the error variables, which

Table 3

Frequently Recommended Goodness-of-Fit Statistics for Confirmatory Factor Analysis Adapted in Part From Byrne (2001)

Statistic group	Statistic[a]	Usual range	Standard for a good fit
Chi-square-based	χ^2	n/a	
	χ^2/df	0 >	2–5
Basic goodness-of-fit measures	RMR	0–1	< .05
	GFI	0–1	.90+
	AGFI	0–1	.90+
Comparisons to baseline model	NFI (BBI)	0–1	.90–.95+
	CFI	0–1	.90–.95+
	IFI	0–1	.90–.95+
Parsimony adjusted	PGFI	0–1	.50+
	PNFI	0–1	.50+
	PCFI	0–1	.50+
Error of approximation	RMSEA	0–1	<. 05–.10
Other indices	AIC	n/a	Relative
	CAIC	n/a	Relative
	ECVI	n/a	Relative

Note. AIC, CAIC, and ECVI are used to compare models, and no absolute standard exists for a good versus poor model fit. Smaller values indicate better fit.
[a]RMR: root mean square residual; GFI: goodness-of-fit index; AGFI: adjusted GFI; NFI: normed fit index (i.e., Bentler-Bonett index); CFI: comparative fit index; IFI: incremental fit index; PGFI: parsimony-adjusted GFI; PNFI: parsimony-adjusted NFI; PCFI: parsimony-adjusted CFI; RMSEA: root mean square error of approximation; AIC: Akaike information criterion; CAIC: consistent AIC; ECVI: expected cross-validation index.

would be based on the belief that several of the self-concept items shared common sources of error. If I believed that to be the case, the model would be reanalyzed with correlated error terms. In this case, correlating three sets of error terms did not result in an appreciable increase in model fit, so the error correlations were not added to the model. The use of modification indexes is probably the most con-

Table 4

Goodness-of-Fit Statistics for Models Using Confirmatory Factor Analysis

Data ECVI	Model	χ^2	*df*	χ^2/df	GFI	PGFI	CFI	PCFI	RMSEA	AIC	CAIC	
1995 TIP	Independence	2026.68	45	45.04	.40	.33	.00	.00	.34–.37	2047	2095	6.06
	Saturated	.00	0	n/a	1.00	n/a	1.00	n/a	n/a	110	375	.32
	One factor	1091.72	35	31.19	.52	.33	.47	.36	.28–.31	1132	1228	3.35
	Two factor	171.06	35	4.89	.91	.58	.93	.72	.09–.12	211	308	.62
	Two factor correlated	166.18	34	4.89	.91	.56	.93	.71	.09–.12	208	310	.62
1997 TIP	Independence	4525.83	45	100.57	.30	.24	.00	.00	.44–.46	4546	4598	9.15
	Saturated	.00	0	n/a	1.00	n/a	1.00	n/a	n/a	110	397	.22
	Two factor	206.41	35	5.90	.93	.59	.96	.75	.09–.11	246	351	.50
	Two factor correlated	177.72	34	5.23	.93	.58	.97	.73	.08–.11	220	329	.44

Note. GFI: goodness-of-fit index; PGFI: parsimony-adjusted GFI; CFI: comparative fit index; PCFI: parsimony-adjusted CFI; RMSEA: root mean square error of approximation, reported as a 90% confidence interval; AIC: Akaike information criterion; CAIC: consistent AIC; ECVI: expected cross-validation index.

troversial aspect of CFA. Critics reason that the exploratory nature of modifying the model stands in sharp contrast to the theory-based, confirmatory nature of CFA. Although these objections have moderated in recent years, researchers should try to make only those modifications that make theoretical sense to the model being tested.

CFA programs provide both unstandardized and standardized output regarding model parameters. When *reporting and interpreting* these data, standardized parameters are preferable, resulting in factor loadings ranging from -1 to 1 and squared multiple correlations (SMCs; i.e., the equivalent of a communality or variance-accounted-for measure in other multivariate techniques) ranging from 0 to 1. For the self-concept data, loadings and SMCs are presented in Figure 2. With the exception of the third math item, factor loadings are large and SMCs provide evidence that each variable is well accounted for by the two-factor model.

Cameron et al. (1997) used CFA to examine the model fit of two different models of intelligence to the Kaufman Assessment Battery for Children (K-ABC) scores of 197 children referred for a gifted program. The authors used several goodness-of-fit measures, including the chi-square statistic, several GFI-based measures, and RMSEA. The results suggested that a model representing the Horn-Cattell fluid-crystallized theory of intelligence fits the data better than a model based on the Kaufman and Kaufman Simultaneous-Sequential-Achievement model of intelligence.

General Guidelines for Best Practice

The following guidelines are intended to serve as suggestions for best practice in conducting, reporting, and interpreting factor analyses. These guidelines are based upon recommendations by Byrne (2001), Bryant and Yarnold (1995), Tabachnick and Fidell (2001), Pedhazur and Shmelkin (1991), and Thompson and Daniel (1996).

Preparation

In most cases, both EFA and CFA require relatively large sample sizes with a minimal amount of missing data. The application of resampling techniques, discussed below, can help provide evidence of reliability in low-sample-size situations, but large samples are preferred. In a related vein, data gathered with instruments marked by low reliability will produce poorer results due to increased error variance. This is especially evident when using EFA. Regardless of which

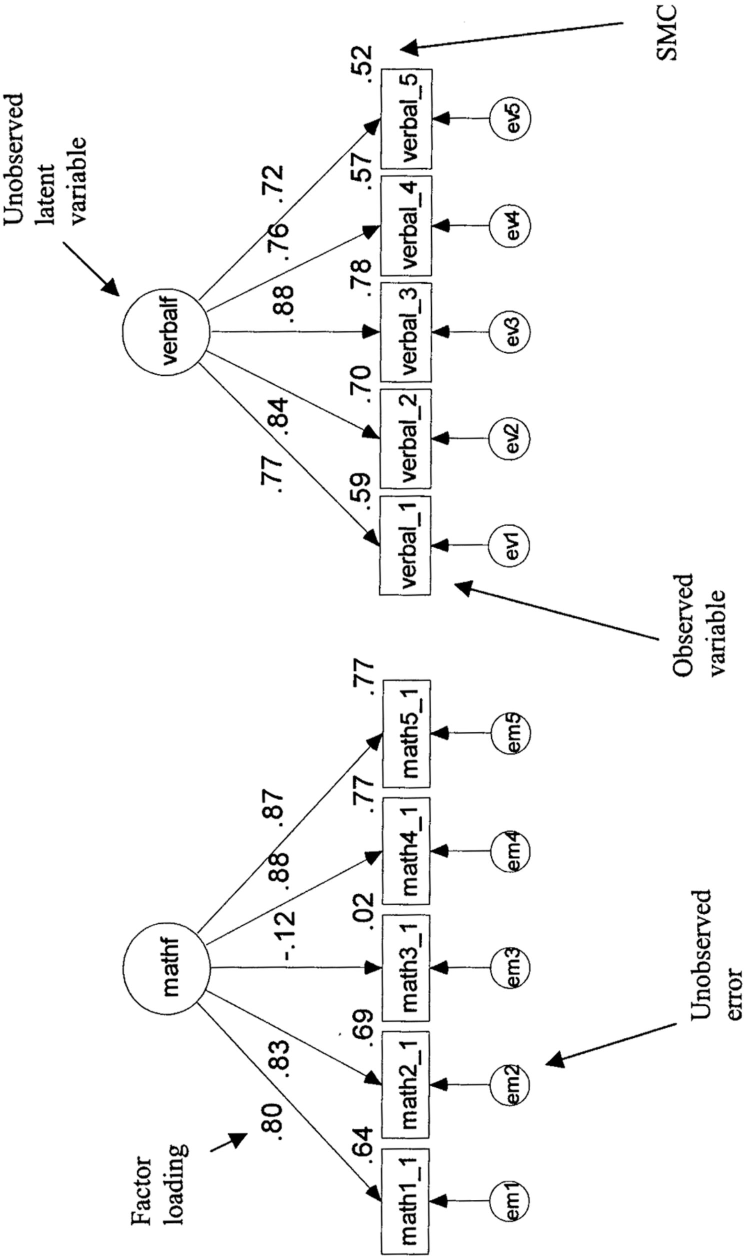

Figure 2. *Standardized parameter estimates for two-factor confirmatory factor analysis model*

technique is used, researchers should briefly explain why they chose EFA, CFA, or a combination of both strategies to analyze data.

Analysis and Reporting

Pedhazur and Schmelkin (1991) provided a list of six necessary pieces of information when reporting the results of EFA: theoretical rationale for use of FA, detailed description of sample and items, methods used (e.g., factor extraction, rotation), criteria employed (e.g., factor selection, loading cutoffs), correlation matrix with means and standard deviations for each variable included in the analyses, structure matrix for orthogonal rotations, and both structure and pattern matrix for oblique rotations (pp. 626–627). These suggestions mirror the recommendations of other researchers for both EFA and CFA reporting. Given the wide variety of techniques for both types of factor analysis, researchers must clearly explain how they conducted their analyses and why specific choices were made.

A necessary step of most factor analyses is the replication of results. In most cases, especially when sample sizes are small or moderate in size, questions emerge about the reliability and replicability of FA results. This is especially true when exploratory methods are used (i.e., EFA) or modification indices are explored (i.e., CFA). The goal of replication/resampling is to determine the extent to which results can be (a) replicated with a new sample that is similar to the original sample or (b) replicated with a large number of randomly selected subsets of the original sample. An excellent example of the utility of such an approach is provided in the study reported by Feldhusen et al. (2000). These researchers randomly selected two subsamples from their relatively small sample of 176 students. Subsequent EFA results were similar to the original analyses using all 176 cases, providing evidence that the initial results were reliable and not negatively influenced by the restricted sample size. Bryant and Yarnold (1995) extended this logic in their recommendation to use EFA and CFA in combination to explore and then confirm factor structure with a sample that can be randomly divided into two groups. Such software programs as AMOS are capable of extensive resampling techniques, such as statistical bootstrapping, to determine the extent to which the results with a particular sample replicate.

In the self-concept example, a second sample of data was collected from students in the same summer program 2 years after the initial data collection. Results from the second round of CFA con-

firmed that the two-factor and two-factor correlated models both have acceptable fit, but the results suggested that the correlated model may provide a slightly better fit. Analysis of model parameters provided evidence that the increased fit may be due to (a) a marginally larger correlation between the factors (-.25 with the replication sample vs. -.18 with the original sample) and (b) the third math variable loading on the math factor consistent with the other math variables. The poor results relative to the third math item in the original analyses may be due to a sample-specific anomaly or reporting or scoring errors. The replication sample was also larger than the first sample (498 vs. 339 students), suggesting a stable solution for a 10-variable model. Without the replication analysis, the construct validity of the math scale could be questioned; with the additional analysis, the results provided evidence that the two academic self-concept scales are associated with acceptable levels of construct validity.

Conclusion

Factor analysis has traditionally been among the most popular multivariate techniques for statistical analysis. However, the technique's longevity often disguises the rapid technical advancements in factor analysis brought about by increased availability of computing power. Many of the weaknesses of factor analysis have been addressed over the past 2 decades, and the widespread availability of structural equation modeling programs has helped make confirmatory factor analyses quite popular in the social and behavioral sciences. Both exploratory and confirmatory techniques are useful tools for analyzing the complex data sets that we frequently encounter when studying giftedness and talent development, and researchers and consumers of research are encouraged to follow guidelines for best practices when conducting and interpreting factor analyses.

References

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: American Psychological Association.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Cameron, L. C., Ittenbach, R. F., McGrew, K. S., Harrison, P. L., Taylor, L. R., & Hwang, Y. R. (1997). Confirmatory factor analysis of the K-ABC with gifted referrals. *Educational and Psychological Measurement, 57*, 823–840.

Carroll, J. B. (1995). Reflections on Stephen Jay Gould's *The Mismeasure of Man* (1981). *Intelligence, 21*, 121–134.

Feldhusen, J. F., Dai, D. Y., & Clinkenbeard, P. R. (2000). Dimensions of competitive and cooperative learning among gifted learners. *Journal for the Education of the Gifted, 23*, 328–342.

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J., & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior, 18*, 667–683.

Kieffer, K. M. (1999). An introductory primer on the appropriate use of exploratory and confirmatory factor analysis. *Research in the Schools, 6*(2), 75–92.

Marsh, H. W. (1992). *Self-Description Questionnaire-II manual*. Cambellton, Australia: University of Western Sydney, Macarthur.

Masten, W. G., Morse, D. T., & Wenglar, K. (1995). Factor structure of the WISC-R for Mexican American students referred for intellectually gifted assessment. *Roeper Review, 18*, 130–131.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Plucker, J. A. (1995). *Measurement and structure of the self-concept of gifted adolescents*. Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Plucker, J. A., & Stocking, V. (2001). Looking outside and inside: Self-concept development of gifted adolescents. *Exceptional Children, 67*, 535–548.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). New York: Allyn and Bacon.

Thompson, B. , & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56*, 197–208.

Author Note

The author appreciates the willingness of the Duke University Talent Identification Program to allow continued access to the data collected from its students. The author also appreciates the constructive criticism provided by Woong Lim and Gayle Dow and the suggestions and continued encouragement of Michael Pyryt.

Endnotes

[1]The data used in the EFA and CFA examples were originally published in Plucker (1995) and Plucker and Stocking (2001).

[2]Statistically, PCA and EFA are quite different: PCA attempts to explain total variance (i.e., common, unique, and error variance), and EFA techniques attempt to explain common variance only.

[3]One advantage of the maximum likelihood method of extraction is the existence of a chi-square test of the number of factors, with statistical insignificance providing evidence that the number of factors is acceptable. In the self-concept data, the chi-square test suggested that additional, substantive factors may exist, $\chi^2(26) = 102.29$, $p < .001$). However, this test is very sensitive to sample size, departure from normality, and other considerations, producing significant results in many if not most situations.